Article

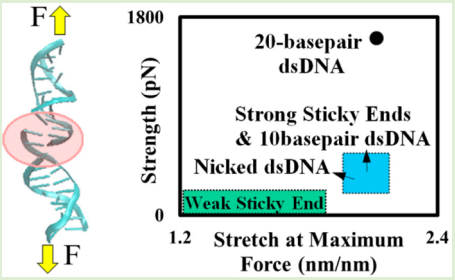
pubs.acs.org/Biomac

Strength of DNA Sticky End Links

Ehsan Ban and Catalin R. Picu*

Department of Mechanical, Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, New York, United States

ABSTRACT: Sticky ends are unpaired nucleotides at the ends of DNA molecules that can associate to link DNA segments. Self-assembly of DNA molecules via sticky ends is currently used to grow DNA structures with desired architectures. The sticky end links are the weakest parts of such structures. In this work, the strength of sticky end links is studied by computational means. The number of basepairs in the sticky end and the sequence are varied, and the response to tension along the axis of the molecule is evaluated using a full atomistic model. It is observed that, generally, increasing the number of basepairs in the sticky end increases the strength, but the central factor controlling this parameter is the basepair sequence. The sticky ends are divided into two classes of low and high strength. The second class has strength comparable with that of a double stranded molecule with one nick in one of the strands. The strength of the first class is roughly half that of the strong sticky ends. For all strong sticky ends tested, the enhanced stability is associated with the formation of an unusually stable complex composed from two basepairs and two flanking bases of certain sequence. This complex rotates and aligns with the direction of the force allowing significant deformation and providing enhanced strength. This is similar to a mechanism recently suggested to enhance the mechanical stability of an RNA kissing loop from the Moloney murine leukemia virus. The model is tested against experimental structural data for sticky ends and against published simulation results for the stretch of double stranded DNA. The results provide guidance for the design of DNA self-assembled structures and indicate the types of sticky ends desirable if maximizing the strength and stability of these structures is targeted.



INTRODUCTION

Single strands of DNA can self-assemble to form artificial structures of desired architectures. Watson−Crick pairing of complementary bases and the ability to synthesize arbitrary sequences of DNA make DNA strands programmable building blocks for constructing a variety of self-assembled structures.[1,2] This includes noncrystalline structures as well as two- and three-dimensional crystals. Examples from the noncrystalline category are provided by structures made by folding longer strands of DNA using smaller auxiliary strands.[3,4] Crystalline structures are made from many smaller molecular segments, such as DNA double-crossovers[5] and tensegrity triangles.[6,7] Further, DNA hybridization can be used to mediate self-assembly of other materials, such as in colloidal crystals.[8,9]

The synthesized two- and three-dimensional crystals have applications in nanoelectronics. They can be used to mediate the self-assembly of nanoscale components such as carbon nanotubes[10] and metallic nanoparticles,[11] which may play the role of electronic components. Also, such structures can be designed to have cavities as large as several hundred nm^3, which can be used to trap macromolecules for structural characterization or as molecular sieves.[12] DNA nanostructures designed using these principles have also been used for targeted drug delivery.[13,14]

DNA crystal structures are made by the self-assembly of DNA building blocks, which are connected through sticky end links. Sticky ends are unpaired bases at the end of a DNA molecule that can associate with complementary bases from a different molecule to provide a link between the two.[15] For example, the 3D DNA crystals synthesized by Zheng et al. are made from identical triangular units. Each triangle is composed from double-stranded DNA (dsDNA) helical domains on the three sides and is connected to six other triangles using sticky ends located at its corners.[7] The three helical domains of a triangle are inclined relative to the triangle plane, which results in a three-dimensional structure with a rhombohedral unit cell.

Free energy calculations based on the nearest-neighbor approximation have been used to design DNA self-assembled structures. The free energy of a desired sequence is calculated as the sum of the free energies for all its dinucleotides. The dinucleotide parameters used in this method come from melting curves of dsDNA molecules.[16] However, it has been shown that other factors, such as the presence of nicks, may affect the stability of these structures.[17] Furthermore, the mechanics of dsDNA may also play a role. Kim et al. used mechanics models to quantitatively predict the shape and flexibility of DNA nanostructures.[18]

Previous studies employing established mechanics methods to investigate biological systems have provided further insight into these issues. Examples include the fracture mechanics-based treatment of the strength of hydrogen-bonded assemblies[19,20] and β-helical protein structures,[21] and studies of DNA packaging using the theory of elasticity.[22,23] Understanding the mechanics of nanoscale building blocks leads to

Received: September 24, 2013
Revised: November 11, 2013
Published: December 6, 2013

   dx.doi.org/10.1021/bm401425k | *Biomacromolecules* 2014, 15, 143−149

improvements of the fidelity of DNA self-assembly models and to design rules for these structures. The flexibility and curvature of pristine dsDNA have been studied extensively both experimentally and by simulations.[24] The stability and rigidity of the DNA double-crossovers has been studied using molecular dynamics simulations.[25,26] Further, Wheatley et al. used the same method to study the dynamics of a DNA Holliday junction.[27] In addition to providing useful data for the design of DNA nanostructures, these studies assist the understanding of molecular mechanics of DNA in vivo. For example, the dynamics of sticky ends and Holliday junctions is relevant for the molecular mechanisms of homologous recombination.

The deformation and the strength of sticky end links with different numbers of basepairs and various sequences are studied in this work using a full atomistic representation and by performing molecular dynamics. Strands of dsDNA with a sticky end link in the central section are subjected to stretch along the axis of the molecule and the force−stretch curve is recorded. Direct observations of the failure mechanism are made. The results lead to a classification of the links as "weak" and "strong," with the strong links containing an unusually stable complex. The effect of the number of basepairs in the sticky end links is also investigated. The results provide guidance for the design of self-assembled DNA structures.

MODEL AND SIMULATION METHOD

An all-atom representation employing the CHARMM27 force field is used.[28,29] The system is advanced in time using the highly scalable NAMD molecular dynamics.[30] B-form DNA structures are generated using the 3D-DART server.[31] The sticky end links are modeled by creating nicks in opposite strands of the BDNA molecules. Then, the structure is solvated using Visual Molecular Dynamics (VMD).[32] The TIP3P model is used for the water molecules.[33] van der Waals interactions are approximated using a 12 Å cutoff. The electrostatic interactions are calculated with the same cutoff. The Ewald summation rule is employed to approximate the long-range electrostatic interactions.[34]

Structures with 10 and 20 basepairs are considered. The simulation cell has dimensions 60 × 60 × 300 Å^3, with the DNA molecule oriented in the longer direction. Periodic boundary conditions are used in all directions. The 20 basepairs molecule has total length of 6.9 nm in the undeformed configuration. The maximum stretch applied is 2.9 nm/nm. This produces molecules up to 20 nm in length. Stretch is calculated by dividing the current length of the molecule by its initial length. Since the tested molecules are much shorter than the persistence length of dsDNA, their contour length is the same as the end-to-end length. In all cases there are at least 12 Å of water separating the nucleic acid atoms and the box boundaries. The equilibrium configuration of a periodic sticky end link with 10 basepairs is also simulated to test the model against structural parameters obtained from X-ray crystallography. In this case a box of height 36 Å is used and the molecule extends across the simulation cell. To make this system periodic, each 3′ end is bonded to the corresponding 5′ end on the other side of the box. In addition, angles and dihedrals that cross the box boundaries in the third dimension are defined.

The systems are charge neutralized by adding K^+ atoms, and then the concentration of KCl is increased to 0.1 M. A solvated DNA molecule is shown in Figure 1. Taking into account water and ions, a typical model contains approximately 100000 atoms.

Energy minimization is performed to eliminate the possible initial excluded volume violations. In this process all DNA atoms are tied to their initial positions using linear springs of constants 500 kcal/mol/Å^2. Time integration is performed in the NVT ensemble and the system is gradually heated up 300 K. Meanwhile, the springs are gradually made softer. This process takes approximately 250 ps. A 2 ns long relaxation period is performed after this step. The equilibration process carried out here is adapted from the method described by Shields et al.[35]

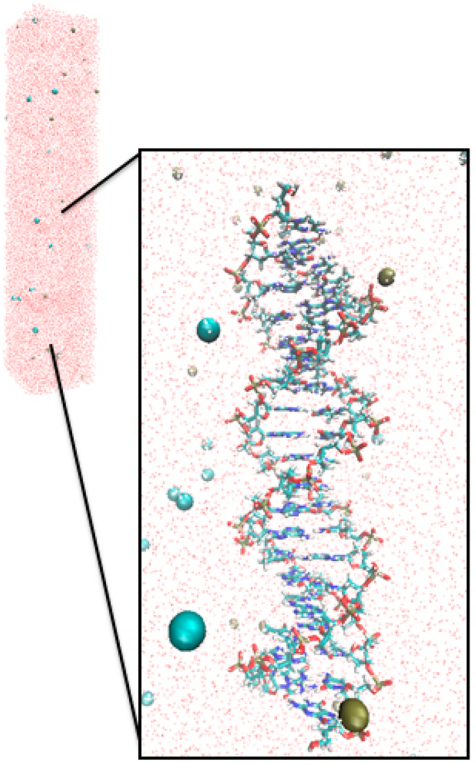


Figure 1. Simulation cell in the initial, undeformed state. The DNA structure is surrounded by water molecules (red and white dots), with a 0.1 M concentration of KCl (K^+ and Cl^- ions are shown as ochre and cyan spheres, respectively).

The molecules are stretched in the NTP ensemble at 300 K and 1 bar after equilibration. The positions of the C5′ atoms at the two ends of the molecule are controlled. One C5′ atom is held fixed, whereas the other one is moved. Position constraints are imposed to these atoms using linear springs of constants 100 kcal/mol/Å^2. In each loading step the moving atom is displaced by 1 Å and then the system is relaxed for 100 ps. The deformation rate is small enough for the water to be in thermodynamic equilibrium with the structure at all times. The forces reported in this work are the averages of the forces in the springs connected to the C5′ atoms performed over the last 20 ps of each deformation step relaxation.

It takes about 10 ns to stretch a molecule up to dissociation using this procedure. The computational effort required is on the order of 1000 processors and 10 h per 1 ns of physical time. Hence, the simulations are quite demanding and performing many replicas of each system is not feasible at this time. We have selected the configurations of interest carefully and have performed replica runs only in few cases.

RESULTS

Tensile Loadings of a dsDNA and Model Validation. To verify the current implementation, the stretch of a dsDNA segment is performed and the results are compared with literature data. This allows observing the failure mode characteristic to this type of loading, which is then used as reference when comparing with the behavior of sticky end links. The model and its implementation are further validated by comparison against experimentally determined structural parameters of sticky ends.

To compare against existing simulations, the stretch of a 20-basepair dsDNA molecule similar to that studied by Severin et

al.[36] is considered. The base sequence is shown in Figure 2a. This structure is equilibrated and loaded, as indicated in section

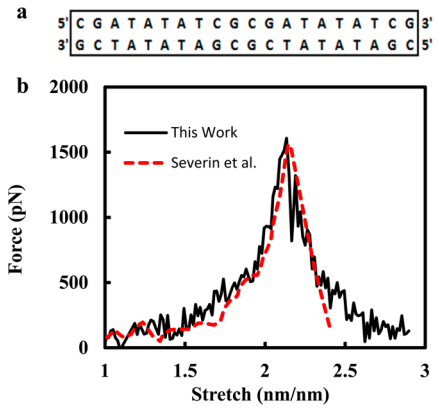


Figure 2. (a) Structure and (b) force against stretch curve for a 20-basepair dsDNA subjected to uniaxial loading. Stretch is calculated by dividing the current length by the initial length. The dashed red line indicates results from the simulation of the same system by Severin et al.[36] The dsDNA molecule has a length of 6.9 nm in the undeformed state.

2. Loading is performed in a slightly different way compared to the work by Severin et al.[36] Specifically, in that work displacements are applied in every time step, while here the system is equilibrated for 100 ps after applying each 1 Å incremental displacement.

The resulting force–stretch curve is shown in Figure 2b. A gradual rise to maximum force is observed, followed by gradual relaxation. The bases at the ends dissociate at a stretch of about 1.9. This is due to the stress concentration at the sites where displacement is applied. Unpeeling from the ends was also observed in optical tweezers experiments by using fluorescence labeling.[37] As the force reaches the peak, the dsDNA transforms into a mix of ladder and zipper-like structures. Balaeff et al. reported similar structures in studying the transition from B-DNA to zip-DNA.[38] Their simulations indicate that zip-DNA has enhanced charge-transport properties compared to B-DNA. After the peak force most of the bases unstack and rotate facing away from the axis of the molecule. Then the two strands slide relative to each other and the measured force decreases.

The force–stretch curves from our work and that of Severin et al.[36] overlap, which provides verification of the current implementation. Unloading is more abrupt in the data from Severin et al.[36] This is due to the different loading scheme used. In our simulations the structure is allowed to relax more in each loading step, which leads to more gradual unloading.

A further test of the model and its implementation is performed by comparing with the sticky end links structural data obtained experimentally by Qiu et al.[15] These structures have a self-complementary sequence with 10 basepairs and two nicks in the paired DNA strands. The two nicks are separated along the axis of the molecule by two basepairs. The sequence is shown in Figure 3a. The twist and rise per base step were obtained by X-ray analysis and are reported by Qiu et al.[15] Here, a periodic self-complementary molecule of the same sequence is simulated. After 12 ns of relaxation, rise and twist angles are measured for each dinucleotide step. The dinucleotide steps are numbered as shown in Figure 3a. These structural parameters fluctuate in the process of relaxation. The values reported are calculated by averaging

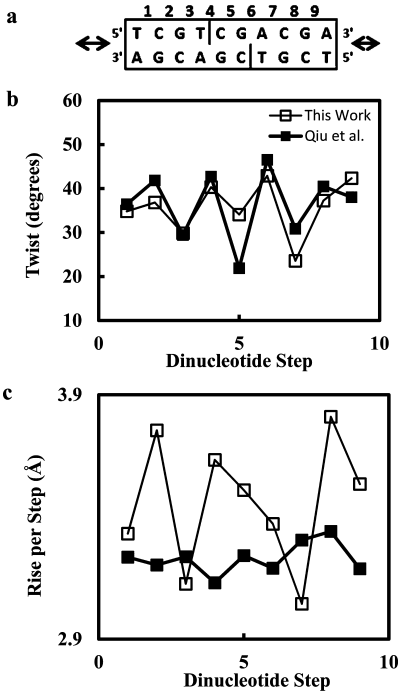


Figure 3. Comparison of structural parameters of a sticky end link of length 2 (open symbols) with experimental data from Qui et al.[15] (filled symbols). (a) The sequence of the periodic sticky end link considered. (b) Twist and (c) rise per step against dinucleotide position. The horizontal axis in (b) and (c) represents the position along the strand, as indicated in (a).

over the final 2 ns of each equilibration step. The calculated parameters are shown in Figure 3b,c along with the experimental data from Qiu et al.[15] As shown in Figure 3b,c, the calculated twist and rise per dinucleotide step agree reasonably well with the corresponding experimental values. Also, the twist value follows the experimental trend.

Stretch of Sticky End Links. Sticky end links of different lengths and base sequences are tested in tension. Figure 4 shows a list of all structures considered. The structure shown in Figure 2a is taken as reference (number **2** in the list of Figure 4). This is a perfect dsDNA, with no nicks, and with an inverted repeat sequence. Its force–stretch curve is shown in Figure 2b and its failure modes are described in the previous section. All other structures tested are derived from this configuration. Structure **3** is identical to **2**, except that it has a nick in one of the strands. The nick separates the dsDNA into two 10-basepair segments. Structure **4** is also identical to **2** and **3**, except that a two basepairs long sticky end link is defined in the middle. The nicks in this structure are separated by two basepairs. To clarify the nomenclature used below, this link is considered to have a length of two.

A total of four structures with two-basepair links are considered: two with the link having CG pairs (numbers **4** and **5**), and two with the link having AT pairs (numbers **6** and **7**). In each group the basepairs adjacent to the link are either CG or AT. This allows for determining the role of the sequence within the link region and the role of the sequence in the immediate vicinity of the link. Structures **8** and **11** have links of CG type, with AT pairs adjacent to them, but with different numbers of bases in the link. Structure **8** has a four-basepair link, whereas **11** has a six-basepair link. Therefore, by comparing structures **5**, **8**, and **11**, the effect of the link length

dx.doi.org/10.1021/bm401425k | Biomacromolecules 2014, 15, 143–149

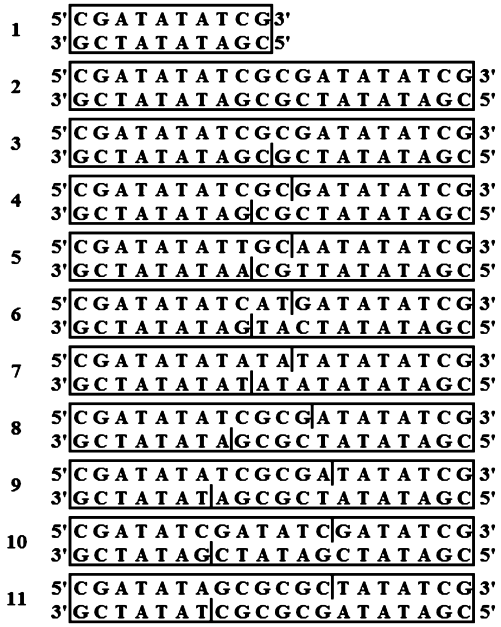


Figure 4. Sequences of the DNA structures considered in this study. Thick vertical lines show the position of nicks in the DNA backbone. Structures **1** and **2** are dsDNA molecules. Structure **3** has only one nick. All other structures have two nicks.

on the strength is determined. Structures **9** and **10** have six-basepair links of different sequences. Finally, stretch of a pristine 10-basepair dsDNA is also simulated (number **1**). Its response is to be compared with that of the nicked molecule. Note that the sequence of structure **1** is different from that shown in Figure 3a. Structure **1** is selected based on the logic discussed here, which unifies all configurations in Figure 4.

The mechanical behavior of these structures is characterized by their force–stretch curves. The parameters considered representative for the mechanical function performed by the sticky end links in the DNA crystals are the maximum force carried before failure, that is, the strength of the link and the stretch corresponding to the maximum force. It turns out that the structures in Figure 4 can be divided in two groups based on their strength: weak and strong.

Dissociation of a weak sticky end-link is a stochastic process that happens over a broad range of forces and an even broader range of stretches. The external force fluctuates significantly before dissociation, but reaches a larger value at the onset of dissociation. Because of the small dissociation force, the two segments connected by the link are not significantly deformed or damaged after dissociation. The force–stretch curve and snapshots of structure **4** (Figure 4) are shown in Figure 5a. Distortion of the link begins with base dissociation at the constrained ends. This isolated event is indicated by the arrow in Figure 5b, which shows the variation of the number of dissociated basepairs during deformation. The process is similar to that observed when stretching the perfect dsDNA, structure **2**. The presence of breaks in the backbone enables the structure to rotate and to release the torsion induced by the stretch. The two segments behave as separate double strands that rotate independently during stretch. Finally, at the maximum stretch, the remaining two basepairs holding the strands together dissociate and the average force drops to zero. This is indicated by the red circle in Figure 5b.

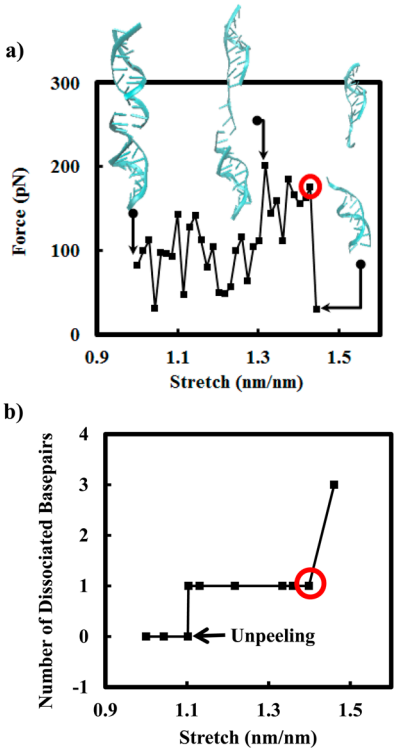


Figure 5. (a) Snapshots and force and (b) number of dissociated basepairs against stretch for a weak sticky end link (structure 4). The red circle in panel (b) marks the onset of dissociation and the black arrow marks unpeeling from a constrained end. The length of the structure in the undeformed state is 6.9 nm.

The strong links dissociate at higher forces and stretches than the weak links. Figure 6 shows the force against stretch, snapshots of the structure, and the number of dissociated bases against stretch for one of the strong sticky end links, structure **8**. The deformation of strong links also begins with base dissociation at the constrained ends. This is marked by arrow 1 in Figure 6b. At later stages of deformation, the bases unstack near the nicks and the structure splits into three domains comprising two double stranded regions at the top and bottom and a complex of six bases in the middle. This is marked by arrow 2 in Figure 6b. The central region includes two basepairs and one flanking base on each side. The flanking bases stabilize the two central basepairs. The deformation continues as the middle domain is holding the two parts together. In the final stage of deformation the two basepairs in the middle region dissociate and the link fails. This is marked by a red circle in Figure 6. Due to the large forces applied, the two segments that separate after link failure are highly distorted. This is marked by the shaded region in Figure 6b. At this stage the structure is distorted and exhibits ladder or zipper-like features that may also include bubbles. These configurations are also observed when stretching a pristine dsDNA.[38,39]

Structures **8** and **11** are examples of strong sticky ends. These links generally dissociate at forces above 400 pN and stretches larger than 2.0 nm/nm. Two replicas of the stretch of structure **8**, denoted as **8** and **8′**, were run. They have similar behavior.

It is useful to compare the force–stretch curves of the perfect dsDNA, structure **2**, with those corresponding to a weak and a strong sticky end link, structures **4** and **8**, respectively. These curves are shown in Figure 7. The strong sticky end link

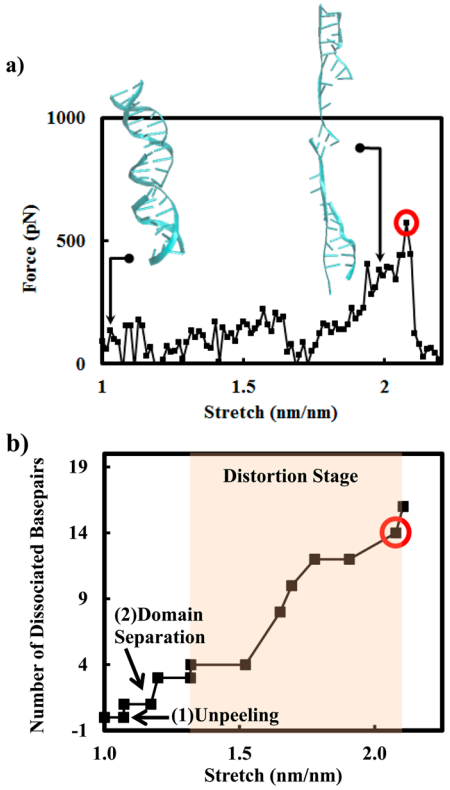


Figure 6. (a) Force–stretch curve along with snapshots of the structure at different stages of deformation and (b) number of dissociated basepairs for a strong sticky end link, structure **8**. In panel (b) arrow 1 indicates unpeeling from a constrained end, arrow 2 indicates the separation of the structure into three helical domains, and the red circle indicates the final step in which the basepairs in the sticky end dissociate. The structure length in the undeformed state is 6.9 nm.

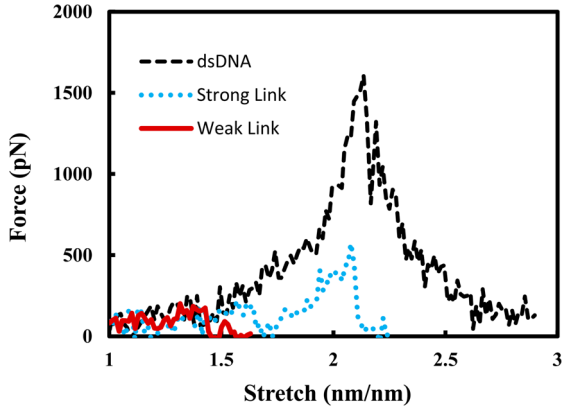


Figure 7. Direct comparisons of force–stretch curves for the pristine dsDNA and weak and strong sticky end links, structures **2**, **4**, and **8**, respectively.

described in Figure 7 has two nicks separated by four basepairs. The structure dissociates at the same stretch as the pristine dsDNA but at a maximum force, which is about 1/3 of the maximum force carried by the dsDNA. This decrease of the structural stiffness in the presence of nicks is discussed in the next section. Structure **4** has two GC base pairs in the sticky end and GC flanking base pairs. This weak link, as well as all studied weak links, does not exhibit the peak, which is characteristic for the dsDNA and the strong links. The failure process of weak links is more stochastic and strongly influenced by thermal fluctuations.

Figure 8 shows the strength and critical stretch of all structures in Figure 4. The cases labeled with a prime

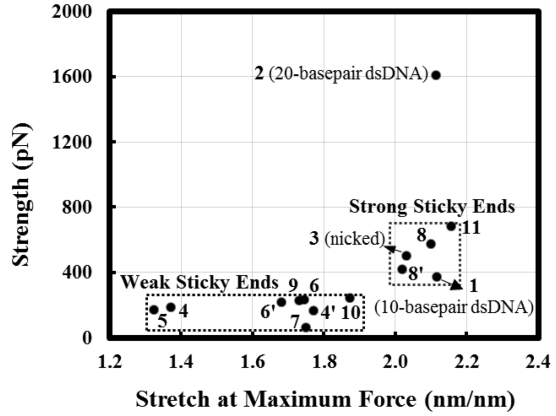


Figure 8. Strength vs stretch at maximum force for all structures listed in Figure 4. The numbers next to the data points correspond to Figure 4. The structures marked with a prime are repeats of the respective configuration. The sticky end links are classified as either weak or strong. The length of all 20-basepair molecules in the underformed configuration is 6.9 nm. The length of the 10-basepair dsDNA is taken to be half this value.

correspond to replicas of the respective structures. The map clearly shows the distinction between strong and weak links. Tests 8, 8′, and 11 correspond to the strong links. These have strengths between 300 and 700 pN and dissociate at stretches larger than 2.0 nm/nm. They are comparable with a dsDNA having only one nick, structure **3**. The other structures are weak, with maximum forces below 250 pN. These fail at stretches smaller than those corresponding to the strong links.

It is useful to inquire what makes a strong link strong. To answer this question one may follow the failure mechanism of a strong link. As described above, a third domain is observed at the onset of dissociation. This segment forms by unstacking of bases close to the two nicks. In all strong links considered, the domain includes two GC basepairs and one flanking guanine on each side. This complex is shown in Figure 9. The planes of all six bases are initially perpendicular to the loading direction and remain perpendicular even at a stretch of 1.0 nm/nm. As deformation progresses, they tend to align with the loading direction. Figure 9d shows the configuration at dissociation (indicated by the red circles in Figure 6). Throughout this process, the two flanking bases stabilize the complex which holds the two DNA segments together. A similar complex was recently reported to enhance the mechanical stability of an RNA kissing loop from the Moloney murine leukemia virus.[40]

DISCUSSION

The selection of structures considered in this work (Figure 4) allows the evaluation of the effect of the sticky end length and sequence on its strength. Let us consider first the effect of the link length. To this end, one compares structures **5**, **8**, and **11**, which have the same type of bases in the link (GC throughout) and same sequence outside the link, but different link lengths. Specifically, 2, 4, and 6 GC basepairs are present in the links of structures **5**, **8**, and **11**, respectively. The strength increases with increasing the link length (Figure 8).

dx.doi.org/10.1021/bm401425k | *Biomacromolecules* 2014, 15, 143–149

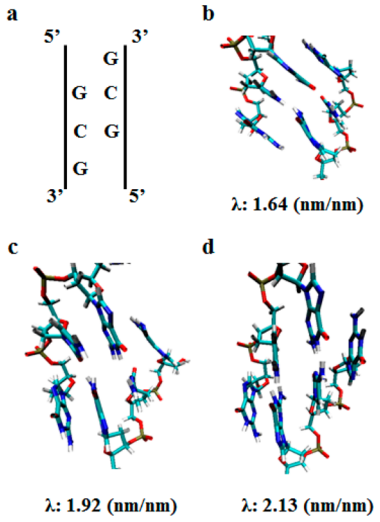


Figure 9. Base sequence and snapshots of the complex stabilizing the strong sticky end links (from the deformation of structure **8**). The stretch, λ, corresponding to each snapshot is indicated. These stretch values correspond to the entire 20-basepair structure as in Figures 5−7

The results indicate that the base sequence plays an important role in defining the strength of the links. The sequence of bases inside the sticky end and the bases adjacent to the link are important. To determine the effect of the sequence within the link one compares structures **9**, **10**, and **11**. Their length is 6 and, except the neighboring basepairs, have the same sequence outside the link. The sequence within the link is different. The strength of structures **9** and **10** is similar, both belonging to the weak class. Structure **11** belongs to the strong category. This indicates that selecting the right sequence leads to a significant increase of the strength. In particular, the large strength is associated with the specific complex described in Results.

Let us observe that for the sequence to matter and for the stabilizing complex to form, the length of the link must be larger than two basepairs. Structure **5**, which has the same sequence as **8** and **11**, but a length of only two basepairs, is weak, while **8** and **11** are strong. In fact, it appears that the sequence is entirely inconsequential in the case of the two-basepair links. All two-basepair links, structures **4**−**7**, are weak.

It is interesting to observe the strong effect the presence of a nick has on the strength of the structure. To this end let us compare the strength of structures **2** and **3**, which have the same length and sequence, except that **3** has a nick in one of the strands. The strength of **3** is about 1/3 the strength of **2**. In fact, the presence of a nick reduces the strength to the range of the strong sticky end links. This happens because at late stages of deformation the presence of nicks permits unstacking of bases at the site of backbone breaks. This results in a weaker coupling between stretch and twist that reduces the apparent stiffness of the structure. Consequently, failure may initiate at the nick, rather than at the strand ends as in the case of a dsDNA.

All simulations are performed with a KCl concentration of 0.1 M. This value is chosen to represent conditions encountered during self-assembling of DNA crystal structures. As well-known, ionic localization may occur in presence of dsDNA, although no such effect is observed in the present simulations. Also, binding of ions is not required for the unusual stability observed in the strong sticky end links and for the formation of the stabilizing complex. However, the maximum force may change with the ionic concentration since experiments indicate that the dsDNA persistence length and stiffness vary when the ionic concentration changes.[41] Similar correlations have been reported in the dynamics of viral DNA ejection,[42] as well as other polyelectrolytes.[43,44]

In addition, the deformation rate has a strong effect on the mechanical behavior of most materials. The effect is usually pronounced in molecular dynamics simulations, in which the deformation rates are usually very large. In the present simulations the effective imposed velocity of one end of the molecule relative to the other is 1 nm/ns. To test the effect of this parameter, additional simulations of the stretch of structure **2** have been performed with a velocity twice as large and at the same temperature (300 K). The resulting force−stretch curve is similar to that in Figure 2, but the maximum force is ∼60% larger. Reducing the test temperature has a similar effect on the force−stretch curve. When structure **2** is deformed at 1 nm/ns and a temperature of 288 K, a curve of shape and maximum force similar to that obtained at the higher rate results.

The forces observed in these simulations are expected to be somewhat larger than the values measured in optical tweezers experiments, primarily due to the lower deformation velocities and longer molecules used in the experiments. To the best of the authors' knowledge, there have been no reports of stretching DNA structures containing sticky end links. The forces required to stretch pristine dsDNA using optical tweezers[45] are orders of magnitude smaller than those resulting from simulations, for example, the response of the 20-basepair dsDNA shown in Figure 2.

Fortunately, these limitations have a smaller bearing on the simulation results reported here for the sticky end links. This is due to the fact that failure in these structures is controlled by processes taking place in the close vicinity of the links. Since the central objective of this work is to determine the strength of the links, the shape of the force−stretch curve is of smaller importance and not capturing the entropic component of the force corresponding to longer time scale conformational changes is not critical. The central result of this work is the ranking of the sticky end links in terms of their strength.

The results presented here can be used in the design of self-assembled DNA structures. A model used for this purpose should take into account the mechanics of dsDNA and single stranded DNA, as well as that of Holliday junctions and of other nanoscale features of the self-assembled structure. The present results provide one of the missing data sets in such models. A specific base sequence is introduced that may be used to stabilize the sticky end links. The stabilization mechanism has a strong mechanics component, and hence, the respective base sequence cannot be predicted exclusively using thermodynamic considerations.

CONCLUSIONS

An analysis of the strength and failure mechanisms of DNA sticky end links is presented in this article. The effect of the sticky end length and base sequence on the strength of the link are studied. It is concluded that the links can be divided in two classes having low and high strengths. The strength increases with increasing the sticky end length, but the critical factor is the sequence. Both the base sequence inside and immediately adjacent to the sticky end link influences its strength. An unusually stable cohesive complex is identified that provides enhanced strength to the links. It includes two basepairs and two flanking bases on each side. The structure and the

stabilizing mechanism for this complex are similar to a complex identified to stabilize an RNA kissing loop from the Moloney murine leukemia virus. These observations can be used in the design of self-assembled DNA structures.

AUTHOR INFORMATION

Corresponding Author

*Tel.: (518) 276-2195. E-mail: picuc@rpi.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

All simulations have been performed on the Blue Gene/L at the Rensselaer Polytechnic Institute Center for Computational Nanotechnology Innovation (CCNI). E.B. also expresses gratitude to Prof. N. C. Seeman for discussions on this subject.

REFERENCES

(1) Seeman, N. C. *Nature* **2003**, *421*, 427−431.
(2) Seeman, N. C. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 225−248.
(3) Rothemund, P. W. K. *Nature* **2006**, *440*, 297−302.
(4) Shih, W. M.; Quispe, J. D.; Joyce, G. F. *Nature* **2004**, *427*, 618−621.
(5) Sa-Ardyen, P.; Vologodskii, A. V.; Seeman, N. C. *Biophys. J.* **2003**, *84*, 3829−3837.
(6) Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C. *Nature* **1998**, *394*, 539−544.
(7) Zheng, J.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R.; Constantinou, P. E.; Ginell, S. L.; Mao, C.; Seeman, N. C. *Nature* **2009**, *461*, 74−77.
(8) Alivisatos, A. P.; Johnsson, K. P.; Peng, X.; Wilson, T. E.; Loweth, C. J.; Bruchez, M. P.; Schultz, P. G. *Nature* **1996**, *382*, 609−611.
(9) Mirkin, C. A.; Letsinger, R. L.; Mucic, R. C.; Storhoff, J. J. A. *Nature* **1996**, *382*, 607−609.
(10) Maune, H. T.; Han, S.-P.; Barish, R. D.; Bockrath, M.; Goddard, W. A.; Rothemund, P. W. K.; Winfree, E. *Nat. Nanotechnol.* **2010**, *5*, 61−66.
(11) Le, J. D.; Pinto, Y.; Seeman, N. C.; Musier-forsyth, K.; Taton, T. A.; Kiehl, R. A. *J. Am. Chem. Soc.* **2004**, 2343−2347.
(12) Paukstelis, P. J. *J. Am. Chem. Soc.* **2006**, *128*, 6794−6795.
(13) Keum, J.-W.; Ahn, J.-H.; Bermudez, H. *Small* **2011**, *7*, 3529−3535.
(14) Lee, H.; Lytton-Jean, A. K. R.; Chen, Y.; Love, K. T.; Park, A. I.; Karagiannis, E. D.; Sehgal, A.; Querbes, W.; Zurenko, C. S.; Jayaraman, M.; Peng, C. G.; Charisse, K.; Borodovsky, A.; Manoharan, M.; Donahoe, J. S.; Truelove, J.; Nahrendorf, M.; Langer, R.; Anderson, D. G. *Nat. Nanotechnol.* **2012**, *7*, 389−393.
(15) Qiu, H.; Dewan, J. C.; Seeman, N. C. *J. Mol. Biol.* **1997**, *267*, 881−898.
(16) SantaLucia, J. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 1460−1465.
(17) Greene, D. G.; Keum, J.-W.; Bermudez, H. *Small* **2012**, *8*, 1320−1325.
(18) Kim, D.-N.; Kilchherr, F.; Dietz, H.; Bathe, M. *Nucleic Acids Res.* **2012**, *40*, 2862−2868.
(19) Keten, S.; Buehler, M. J. *Phys. Rev. Lett.* **2008**, *100*, 198301−198304.
(20) Keten, S.; Buehler, M. J. *Nano Lett.* **2008**, *8*, 743−748.
(21) Keten, S.; Buehler, M. J. *Comput. Methods Appl. Mech. Eng.* **2008**, *197*, 3203−3214.
(22) Purohit, P. K.; Kondev, J.; Phillips, R. *J. Mech. Phys. Solids* **2003**, *51*, 2239−2257.
(23) Purohit, P. K.; Kondev, J.; Phillips, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3173−3178.
(24) Peters, J. P.; Maher, L. J. *Q. Rev. Biophys.* **2010**, *43*, 23−63.
(25) Maiti, P. K.; Pascal, T. A.; Vaidehi, N.; Heo, J.; Goddard, W. A. *Biophys. J.* **2006**, *90*, 1463−1479.
(26) Santosh, M.; Maiti, P. K. *Biophys. J.* **2011**, *101*, 1393−1402.
(27) Wheatley, E. G.; Pieniazek, S. N.; Mukerji, I.; Beveridge, D. L. *Biophys. J.* **2012**, *102*, 552−560.
(28) Foloppe, N.; Mackerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86−104.
(29) Mackerell, A. D.; Banavali, N. K. *J. Comput. Chem.* **2000**, *21*, 105−120.
(30) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(31) Van Dijk, M.; Bonvin, A. M. J. J. *Nucleic Acids Res.* **2009**, *37*, W235−W239.
(32) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics Modell.* **1996**, *14*, 33−38.
(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(34) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(35) Shields, G. C.; Laughton, C. a.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7463−7469.
(36) Severin, P. M. D.; Zou, X.; Gaub, H. E.; Schulten, K. *Nucleic Acids Res.* **2011**, *39*, 8740−8751.
(37) Van Mameren, J.; Gross, P.; Farge, G.; Hooijman, P.; Modesti, M.; Falkenberg, M.; Wuite, G. J. L.; Peterman, E. J. G. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18231−18236.
(38) Balaeff, A.; Craig, S. L.; Beratan, D. N. *J. Phys. Chem. A* **2011**, *115*, 9377−9391.
(39) King, G. A.; Gross, P.; Bockelmann, U.; Modesti, M.; Wuite, G. J. L.; Peterman, E. J. G. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 3859−3864.
(40) Chen, A. A.; García, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E1530−E1539.
(41) Smith, S. B.; Finzi, L.; Bustamante, C. *Science* **1992**, *258*, 1122−1126.
(42) Wu, D.; Van Valen, D.; Hu, Q.; Phillips, R. *Biophys. J.* **2010**, *99*, 1101−1109.
(43) Cranford, S. W.; Buehler, M. J. *Soft Matter* **2013**, *9*, 1076−1090.
(44) Cranford, S. W.; Ortiz, C.; Buehler, M. J. *Soft Matter* **2010**, *6*, 4175−4185.
(45) Bustamante, C.; Smith, S. B.; Liphardt, J.; Smith, D. *Curr. Opin. Struct. Biol.* **2000**, *10*, 279−285.